United States Patent [19]

Roberts et al.

[11] **Patent Number: 5,027,406**

[45] **Date of Patent: Jun. 25, 1991**

[54] **METHOD FOR INTERACTIVE SPEECH RECOGNITION AND TRAINING**

[75] Inventors: **Jed Roberts**, Cambridge; **James K. Baker**, West Newton; **Edward W. Porter**, Boston, all of Mass.

[73] Assignee: **Dragon Systems, Inc.**, Newton, Mass.

[21] Appl. No.: **280,700**

[22] Filed: **Dec. 6, 1988**

[51] **Int. Cl.**[5] **G10L 7/08**

[52] **U.S. Cl.** **381/43**

[58] **Field of Search** 381/41-46, 381/110; 364/513.5

[56] **References Cited**

U.S. PATENT DOCUMENTS

4,277,644	7/1981	Levinson et al.	381/43
4,653,097	3/1987	Watanabe et al.	381/42
4,817,156	3/1989	Bahl et al.	381/43
4,829,576	5/1989	Porter	381/43
4,829,577	5/1989	Kuroda et al.	381/43
4,866,778	9/1989	Baker	381/43
4,882,757	11/1989	Fisher et al.	381/43

FOREIGN PATENT DOCUMENTS

0002039 1/1981 Japan .

OTHER PUBLICATIONS

Nassimbene, "Speech Technique for Data Entry to Word Processing Equipment", IBM Technical Disclosure, vol. 22, No. 7, 12/79, pp. 2891-2892.

Primary Examiner—Dale M. Shaw
Assistant Examiner—John A. Merecki
Attorney, Agent, or Firm—Edward W. Porter

[57] **ABSTRACT**

A method for creating word models for a large vocabulary, natural language dictation system. A user with limited typing skills can create documents with little or no advance training of word models. As the user is dictating, the user speaks a word which may or may not already be in the active vocabulary. The system displays a list of the words in the active vocabulary which best match the spoken word. By keyboard or voice command, the user may choose the correct word from the list or may choose to edit a similar word if the correct word is not on the list. Alternately, the user may type or speak the initial letters of the word. Then the recognition algorithm is called again satisfying the initial letters, and the choices displayed again. A word list is then also displayed from a large backup vocabulary. The best words to display from the backup vocabulary are chosen using a statistical language model and optionally word models derived from a phonemic dictionary. When the correct word is chosen by the user, the speech sample is used to create or update an acoustic model for the word, without further intervention by the user. As the system is used, it also constantly updates its statistical language model. The system gets more and more word models and keeps improving its performance the more it is used. The system may be used for connected speech as well as for discrete utterances.

48 Claims, 21 Drawing Sheets